AD-A083 782 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 12/1

ON AN EXPONENTIAL SERVER WITH GENERAL CYCLIC ARRIVALS.(U)

FEB 80 A K AGRAWALA, S K TRIPATHI AFOSR-78-3654

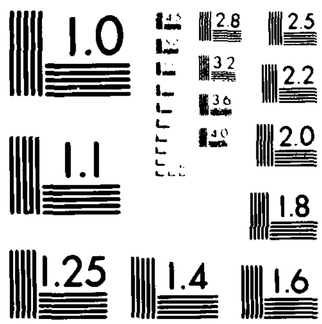
UNCLASSIFIED CSC-TR-867 AFOSR-TR-80-0311 NL

END

DATE FILMED

6-80

DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

12

ADA083782

Technical Report TR-867 February 1980

[redacted]

ON AN EXPONENTIAL SERVER
WITH GENERAL CYCLIC ARRIVALS(1)

by

Ashok K. Agrawala and Satish K. Tripathi

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER
COLLEGE PARK, MARYLAND

Technical Report TR-867 February 1980

[redacted]

ON AN EXPONENTIAL SERVER WITH GENERAL CYCLIC ARRIVALS(1)

by

Ashok K. Agrawala and Satish K. Tripathi

(1) This research was supported in part by the Air Force Office of Scientific Research under grant AFOSR 78-3654.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 19 AFOSR-TR-80-0311
2. GOVT ACCESSION NO.: AD-A083 782
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): 6 ON AN EXPONENTIAL SERVER WITH GENERAL CYCLIC ARRIVALS
5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): 10 Ashok K. Agrawala, Satish K. Tripathi
8. CONTRACT OR GRANT NUMBER(s): 15 AFOSR-78-3654

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
University of Maryland
Computer Science Center
College Park, MD 20742

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 16 17 61102F 2304 A2

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Office of Scientific Research/NM
Bolling AFB, Washington, DC 20332

12. REPORT DATE: 11 Feb [redacted] 80
13. NUMBER OF PAGES: 33

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 9 Technical rept.,
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

12 25

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

14 CSC-TR-867

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

queueing theory, expected virtual waiting time, cyclic arrivals, deterministic routing, spectoral factorization

(superscript n)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A $G^{n}/M/1$ queue is defined as a single server queue with exponential service time and general cyclic arrival distributions of cycle length n. The waiting time distribution for such a queue is proved to be a sum of n exponential terms; this is a generalization of g/M/1 queue results. Based on this a method for obtaining the steady-state waiting time distributions for $G^{n}/M/1$ queues is introduced. An example is presented to show an application of $G^{n}/M/1$ queues in deterministic routing.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 403018

ON AN EXPONENTIAL SERVER WITH GENERAL CYCLIC ARRIVALS[1]

Ashok K. Agrawala and Satish K. Tripathi
Department of Computer Science
University of Maryland
College Park, MD

ABSTRACT

A $G^n|M|1$ queue is defined as a single server queue with exponential service time and general cyclic arrival distributions of cycle length n. The waiting time distribution for such a queue is proved to be a sum of n exponential terms; this is a generalization of $G|M|1$ queue results. Based on this a method for obtaining the steady-state waiting time distributions for $G^n|M|1$ queues is introduced. An example is presented to show an application of $G^n|M|1$ queues in deterministic routing.

Key words and phrases: queueing theory, expected virtual waiting time, cyclic arrivals, deterministic routing, spectoral factorization.

CR Categories: 3.8, 5.5, 8.1

(1) The work was supported in part by the U.S. Air Force Office of Scientific Research under grant AFOSR 78-3654 to the Department of Computer Science.

I. INTRODUCTION

In this paper we consider the behavior of single server queues in which the service time distribution S(t) is exponential. The i^{th} interarrival time is assumed to have a distribution $A_i(t)$ (or density $a_i(t)$). In the most general case all A_i's may be different. Here, however, we consider the case where the sequence of A_i's consists of a repeating cycle of length n. In other words the k^{th}, $n+k^{th}$, $2n+k^{th}$, ... arrivals have the same interarrival distribution A_k for k=1 ... n. These distributions are assumed to have rational Laplace transforms. We refer to this queueing system as $G^n|M|1$ system. Such queueing systems arise in computer networks using deterministic routing.

The behaviour of a $G^n|M|1$ system may be characterized in terms of its virtual waiting time where the virtual waiting time at any instant is the unfinished work at the server. This is a well known approach used in the study of GI|G|1 systems [1, 2, 3, 4].

The virtual waiting time for a G|M|1 queue is known to have an exponential steady state distribution [3, 5]. Here we prove that the steady state distribution for the virtual waiting time of a $G^n|M|1$ queue is a sum of n exponential terms. The approach taken in this paper uses the decay characteristics of the expected virtual waiting time in the future. In section II after introducing the necessary notation, the equivalence of the expected virtual waiting time decay function and the virtual waiting time distribution function are established. The solution for the steady state expected virtual waiting time decay function for a $G^n|M|1$ queuing system is obtained in section III. An example for application of the results of a $G^n|M|1$ queue is given in Section IV.

II. NOTATION

The approach used here is based on the results presented in [4]. For the sake of completeness, in the following we present the notations and results relevant to the discussion here.

Consider a server providing service to arriving customers according to a first-in-first-out discipline. Let the service need of a customer be determined probabilistically and defined by a service time distribution $S(t)$ (or a corresponding service time density function $s(t)$). We are interested in characterizing this queuing system in terms of the virtual waiting time $\tau(t)$ at time instant t. Let the density and the distribution functions of $\tau(t)$ be $w(\tau,t)$ and $W(\tau,t)$, respectively. Clearly, these functions are defined for $\tau \geq 0$ and $t \geq 0$.

In the virtual waiting time density $w(\tau,t)$, τ is the random variable and t is a parameter of this density function. For a given value of t, τ specifies the waiting time from t onwards and hence t specifies the origin for τ. When we change the value of τ to t_1 in $w(\tau,t)$ the origin for τ is changed with it to reflect the waiting time corresponding to the instant t_1.

When an arrival occurs at time T the waiting time density function immediately after the arrival, $w(\tau,T^+)$ may be expressed as

$$w(\tau,T^+) = \int_0^{\tau} w(\sigma,T^-)\ s(\tau-\sigma)d\sigma \qquad \rightarrow (2.1)$$

As a result of the arrival, the waiting time undergoes a step change in that $\tau(T^+) = \tau(T^-)$ + service time. Therefore, the density of the waiting time after the arrival at time T is obtained by convolving $w(\tau,T^-)$ with the density function of the service time $s(\tau)$. Similarly, the distribution function of the waiting time after an arrival at time T is given as

$$W(\tau,T^+) = \int_0^{\tau} W(\sigma,T^-)\ s(\tau-\sigma)d\sigma \qquad \rightarrow (2.2)$$

or

$$W(\tau,T^+) = W(\tau,T^-) \circledast s(\tau) \qquad \rightarrow (2.3)$$

where $\circledast$ is the convolution operator.

Having seen how $w(\tau,t)$ changes when an arrival occurs, let us see how it changes with time when no arrival occurs. Given $w(\tau,T)$ and that no arrivals occur between T and $T+T_1$, $w(\tau,T+T_1)$ is given as

$$w(\tau,T+T_1) = \delta(\tau)\int_0^{T_1} w(\sigma,T)d\sigma + w(\tau+T_1,T) \qquad \tau \geq 0 \rightarrow (2.4)$$

and

$$W(\tau,T+T_1) = W(\tau+T_1,T) \qquad \tau \geq 0 \rightarrow (2.5)$$

where $\delta(\tau)$ is the unit impulse function [3].

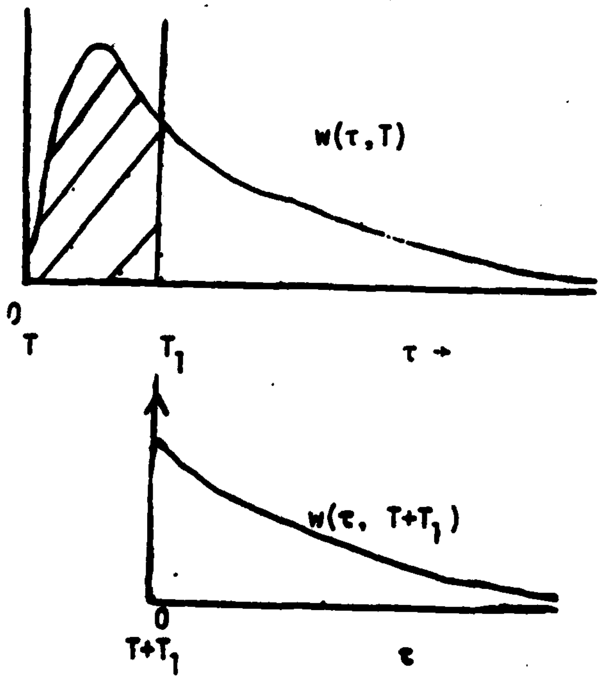


Figure 1

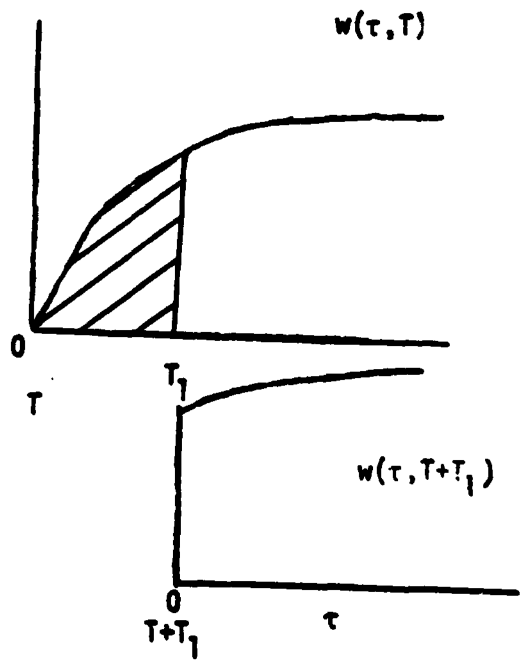


Figure 2

The effect of no arrivals between time T and $T+T_1$ is explained in Figure 1. The origin of τ for $w(\tau, T+T_1)$ is changed to $T+T_1$ as shown in the figure. The magnitude of the delta function at $\tau=0$ for $w(\tau, T+T_1)$ corresponds to the probability of the event $\{0 \leq \tau(T^-) < T_1\}$. This probability is simply the shaded area of $w(\tau,T)$. The corresponding changes to the distribution function (Equation 2.5) are shown in Figure 2.

If we are given the arrival instants and the initial waiting time distribution, we may compute $w(\tau,t)$ for any time t using equations (2.1)-(2.5). Note that no assumptions are made about the characteristics of the arrival process. The complete transient solution for the waiting time distribution may thus be obtained.

If we are only interested in the epochs just before or just after the arrivals and not necessarily at particular time instants we may use the following result:

Let $w_{n+}(\tau,0)$ be the waiting time density function after the n^{th} arrival at time 0. The waiting time density function just before the next arrival n+1 is given as

$$w_{(n+1)^-}(\tau) = \int_0^\infty w_n(\tau,t)\ a_n(t)dt \qquad \rightarrow (2.6)$$

where $a_n(t)$ is the interarrival time density for the next arrival and $w_n(\tau,t)$ is computed from $w_n(\tau,0)$ using equation (2.4).

These results can also be used to obtain the steady state solution, if one exists. To arrive at the steady state behavior we observe that the waiting time distribution remains the same immediately before (or immediately after) each arrival. Let an arrival (say the n^{th}) occur at time 0 and the system is in steady-state. Then, given the next arrival occurs at t, $w(\tau,t)$ is obtained using equation (2.4). But as random arrivals may occur, $w(\tau,t)$ is only a conditional density function for the waiting time seen by the next arrival. This marginal density is given by equation (2.6). Under the steady state condition, we may equate either

$w_{n-}(\tau)$ to $w_{n+1^-}(\tau)$, or $w_{n+}(\tau)$ to $w_{n+1^+}(\tau)$. Therefore, $W(\tau)$ is the steady state waiting time distribution function if and only if $W(\tau)$ satisfies

$$W(\tau) = \int_0^\tau \int_0^\infty W(\sigma+t)\ a(t)\ s(\tau-\sigma)dtd\sigma \qquad \rightarrow (2.7)$$

Note that by rearranging the terms in equation (2.7) Lindley's integral equation [2] results.

While $\tau(t)$ is a stochastic process which is completely characterized by the desity function $w(\tau,t)$, in many practical situations we are interested in calculating the expected value of $\tau(t)$, $E_w(t)$. Knowing $w(\tau,t)$, $E_w(t)$ may be calculated using the following equation

$$E_w(t) = \int_0^\infty \tau\ w(\tau,t)d\tau \qquad \rightarrow (2.8)$$

Note that equation (2.8) involves $w(\tau,t)$. In order to use it in calculating $E_w(t)$ for an arbitrary t, $w(\tau,t)$ must be known for all τ.

We define $E_{w+}(t)$ and $E_{w-}(t)$, to be the expected virtual waiting time at instant t evaluated immediately after and before an arrival occuring at t=0, respectively.

As shown in [4], $E_w(t)$ before and after an arrival is related as follows.

$$E_{w+}(t) = E_s(t) + E_{w-}(0)(1-S(t)) + s(t) \circledast E_{w-}(t) \qquad \rightarrow (2.9)$$

and

$$E_{w-}(t) = \int_0^\infty a(T)\ E_{w+}(t+T)dT \qquad t \geq 0 \qquad \rightarrow (2.10)$$

where $\quad E_s(t) = \int_0^\infty T\ s(t+T)dT. \qquad \rightarrow (2.11)$

Equation (2.9) allows us to compute $E_{w+}(t)$ given $E_{w-}(t)$ and the occurence of an arrival. Equation (2.10) may then be used to evaluate $E_{w-}(t)$ just before the next arrival given the interarrival density function a(t) for the next arrival. These equations can be used to evaluate the steady state values for these functions when such a steady state exists.

The relationship between $E_w(t)$ and $w(t,0)$ is established in Theorem 1.

<u>Theorem 1</u> - $E_w(t)$ and $w(t,0)$ are equivalent in that given one the other is uniquely determined.

<u>Proof</u> - (a) Given $w(t,0)$, equations (2.8) and (2.5) uniquely determine $E_w(t)$.

(b) Given $E_w(t)$ we may write [5]

$$E_w(t) = \int_0^\infty [1-W(\tau,t)]d\tau.$$

From equation (2.5)

$$E_w(t) = \int_0^\infty [1-W(t+\tau, 0)]d\tau$$

$$= \int_t^\infty [1-W(v,0)]dv.$$

Let $1-W(v,0) = U(v,0)$

Thus, $E_w(t) = \int_t^\infty U(v,0)dv.$

Also, $E_w(t+\Delta t) = \int_{t+\Delta t}^\infty U(v,0)dv.$

Therefore, $$\frac{E_w(t+\Delta t) - E_w(t)}{\Delta t} = -\frac{1}{\Delta t}\int_t^{t+\Delta t} U(v,0)dv$$

In the limit as $\Delta t \to 0$

$$\frac{dE_w(t)}{dt} = -U(t,0) = -(1-W(t,0)).$$

Thus $$W(t,0) = \frac{dE_w(t)}{dt} + 1.$$

This completes the proof of Theorem I. ■

The approach taken in this paper uses $E_w(t)$ to characterize the queuing system. Clearly, from $E_w(t)$ the distribution functions for the virtual waiting time can be evaluated.

III. SOLUTION TO $G^n|M|1$ SYSTEM

Consider a first-in-first-out exponential server with

$$s(t) = \mu e^{-\mu t}.$$

Let the i^{th} interarrival distribution and density be $A_i(t)$ and $a_i(t)$, respectively. Also, let $E_{w-_i}(t)$ and $E_{w+_i}(t)$ be the expected virtual waiting time decay functions just before and after the i^{th} arrival.

Assume that the arrivals to this server consist of a repeating cycle of length n in that

$$a_{i+kn}(t) = a_i(t) \qquad \text{for } i = 1,2,\ldots,n \text{ and } k = 1,2,3\ldots.$$

In the steady state, if one exists,

$$\left.\begin{aligned} &E_{w+_{i+kn}}(t) = E_{w+_i}(t) \\ \text{and} \quad &\\ &E_{w-_{i+kn}}(t) = E_{w-_i}(t). \end{aligned}\right\} \qquad \text{for } i = 1,2,\ldots,n \text{ and } k = 1,2,3\ldots. \qquad \rightarrow (3.1)$$

To solve this system we use equation (2.9) to evaluate $E_{w+_i}(t)$ given $E_{w-_i}(t)$ and use equation (2.10) to evaluate $E_{w-_{i+1}}(t)$ given $E_{w+_i}(t)$. These combined with equation (3.1) are then used to obtain the steady state values of $E_w(t)$ functions. Note that for the exponential service time

$$E_s(t) = \frac{1}{\mu} e^{-\mu t}$$

and

$$S(t) = 1 - e^{-\mu t}$$

From equation (2.9)

$$E_{w+_i}(t) = \frac{1}{\mu} e^{-\mu t} + E_{w-_i}(0)(e^{-\mu t}) + E_{w-_i}(t) \circledast \mu e^{-\mu t}$$

or

$$E_{w+_i}(t) = (E_{w-_i}(0) + \frac{1}{\mu})e^{-\mu t} + E_{w-_i}(t) \circledast \mu e^{-\mu t} \qquad t \geq 0$$

Also,

$$E_{w-_i}(0) + \frac{1}{\mu} = E_{w+_i}(0)$$

Thus,

$$E_{w+_i}(t) = E_{w+_i}(0)e^{-\mu t} + E_{w-_i}(t) \circledast \mu e^{-\mu t} \qquad \rightarrow (3.2)$$

$$t \geq 0$$

From equation (2.10)

$$E_{w-_{i+1}}(t) = \int_0^\infty a_{i+1}(T)\ E_{w+_i}(t+T)dT \quad t \geq 0 \qquad \rightarrow (3.3)$$

Let us define the Laplace transforms of $E_{w+_i}(t)$ & $E_{w-_i}(t)$ as

$$L_{w+_i}(s) = \int_0^\infty e^{-st}E_{w+_i}(t)dt \qquad s \geq 0 \qquad \rightarrow (3.4)$$

and
$$L_{w-_i}(s) = \int_0^\infty e^{-st}E_{w-_i}(t)dt \qquad s \geq 0 \qquad \rightarrow (3.5)$$

Note that as $E_{w+_i}(t)$ and $E_{w-_i}(t)$ are defined for $t \geq 0$, $L_{w+_i}(s)$ and $L_{w-_i}(s)$ are analytic in the right halfplane of s. Also, as $E_w(t)$ functions are positive and integrable the transforms $L_w(s)$ exist [3].

Taking the transform of equation (3.2) we may write

$$L_{w+_i}(s) = E_{w+_i}(0)\ \frac{1}{\mu+s} + L_{w-_i}(s) \cdot \frac{\mu}{\mu+s} \qquad \rightarrow (3.6)$$

Equation (3.3) expresses a form of convolution which is defined only for $t \geq 0$. If we were to add a function of the form

$$\int_0^\infty a_{i+1}(T)E_{w+_i}(t+T)dT \qquad t < 0$$

we may write the corresponding tranforms as [3]:

$$G^-_{i+1}(s) + L_{w-_{i+1}}(s) = L_{a_{i+1}}(-s) \cdot L_{w+_i}(s) \qquad \rightarrow (3.7)$$

where $G^-_{i+1}(s)$ is a function of s which is analytic in the left half plane of s.

Before proceeding further we need the following Lemmas.

Let a(t) be the density function of the interarrival time with a rational laplace transform $L_a(s)$.

<u>Lemma 1</u> $\frac{\mu}{\mu+s} L_a(-s)$ can be written as

$$\frac{\mu}{\mu+s} L_a(-s) = \phi^+(s) + \phi^-(s) \qquad \rightarrow (3.8)$$

where

$\phi^+(s)$ is analytic for $Re(s) > 0$ and $\phi^-(s)$ is analytic for $Re(s) < 0$.

And that $\phi^+(s)$ has the form $\phi^+(s) = \frac{c}{\mu+s}$ where c is a constant.

Proof. As $L_a(s)$ is rational we may write it as

$$L_a(s) = \frac{\alpha_{n-1}s^{n-1} + \alpha_{n-\ell}s^{n-\ell} + \ldots + \alpha_0}{\beta_n s^n + \beta_{n-1}s^{n-1} + \ldots + \beta_0}$$

Note that a(t) is defined for $t > 0$ only. Thus, $L_a(s)$ is analytic in the right half plane and has poles only in the left half plane.

Therefore, $L_a(-s)$ is analytic in the left half plane and has poles only in the right half plane. Clearly, the product

$$\frac{\mu}{\mu+s} L_a(-s)$$

has one pole in the left half plane and the rest in the right half plane.

Therefore, this function can be written as $\phi^+(s) + \phi^-(s)$ where

$$\phi^+(s) = \frac{c}{\mu+s} \text{ where c is a constant.}$$

This completes the proof. ■

Lemma 2

$$\left(\frac{\mu}{\mu+s}\right)^n \prod_{i=1}^{n} L_{a_i}(-s) \text{ can be written as } \phi^+(s) + \phi^-(s)$$

where

$$\phi^+(s) = \sum_{i=1}^{n} \frac{c_i}{(\mu+s)^i} \qquad \rightarrow (3.9)$$

Proof. This directly follows from the Lemma 1. ■

We observe that multiplying the expression by an rational function ψ^- which is analytic in the left half plane will still yield a decomposition in which $\Phi^+(s)$ has the form shown in equation (3.9)

<u>Lemma 3</u>

$$L_{w-_n}(s) = x(s)L_{w-_0}(s) + Y(s) - z(s)$$

where

$$x(s) = \left[\prod_{i=1}^{n}\left(\frac{\mu}{\mu+s}\right) L_{a_i}(-s)\right]$$

$$Y(s) = \sum_{i=1}^{n} \mu^{n-i} E_{w+_{i-1}}(0) \left\{\prod_{j=i}^{n} \frac{L_{a_j}(-s)}{\mu+s}\right\}$$

and

$$z(s) = \sum_{i=1}^{n-1} \left\{\prod_{j=i+1}^{n} \frac{\mu}{\mu+s} L_{a_j}(-s)\right\} G^-_i(s) \qquad \to (3.10)$$

<u>Proof</u>. Follows from a repetitive application of equations (3.6) and (3.7). ■

For the next two Lemmas let us consider a distribution function $F(x)$, $x \geq 0$ with expected value β and Laplace transform $\mathcal{L}_f(s)$.

<u>Lemma 4</u> If (a) $Re(s) \geq 0$, $|w| < 1$ or (b) $Re(s) > 0$, $|w| \leq 1$ or (c) $\mu\beta > n$ and $Re(s) \geq 0$, $|w| \leq 1$ then the equation

$$z^n = w\mathcal{L}_f(s + \mu(1-z))$$

has exactly m roots $z = \delta_r(s,w)(r=1,2,\ldots m)$ in the unit circle $|z| < 1$.

We have

$$\delta_r(s,w) = \sum_{j=1}^{\infty} \frac{(-\mu)^{j-1}(\epsilon_r w^{1/m})^j}{j!} \left(\frac{d^{j-1}[\Phi(\mu+s)]^{j/m}}{ds^{j-1}}\right) \qquad \to (3.11a)$$

where $\epsilon_r = e^{2\Pi ir/n}$ $(r=1,2,\ldots m)$ are the m^{th} roots of unity.

<u>Proof</u>. See [6, page 126]. ■

<u>Lemma 5</u> Let $\delta_r = \delta_r(0,1)$ where $\delta_r(s,w)$ is given by equation (3.11a). If $\mu\beta > m$ then $\delta_1,\delta_2\ldots\delta_m$ are the m roots in z of the equation

$$z^m = \mathcal{L}_f(\mu(1-z)) \quad \rightarrow (3.11b)$$

in the unit circle $|z| < 1$. If $\mu\beta \le m$ then $\delta_1, \delta_2, \ldots \delta_{m-1}$ are the m-1 roots in z of (3.11b) in the unit circle $|z| < 1$, whereas $\delta_m = 1$.

Proof. See [6, page 126]. ■

We further observe that equation (3.11b) always has one root at $z = 1$. This is because for $\mathcal{L}_f(s)$ to be a Laplace transform of a probability distribution $\mathcal{L}_f(0) = 1$

Theorem 2 - In a $G^n|M|1$ queue where the arrival distributions have rational Laplace transform,

$$L_{W-}(s) = \sum_{i=1}^{n} \frac{\alpha_i}{s+\zeta_i} \quad \rightarrow (3.12)$$

where $L_{W-}(s)$ is the Laplace transform of $E_{W-}(t)$ at the beginning of a cycle, α_i's are constants and ζ_i's are the n roots of the equation

$$[1 - \prod_{i=1}^{n} \frac{\mu}{\mu+s} L_{a_i}(-s)] = 0$$

in the region $|s| < \mu$.

Proof - For a $G^n|M|1$ queue with a cycle of length n, in the steady state, if one exists, $L_{W-_n}(s) = L_{W-_0}(s) = L_{W-}(s)$. In order to find $L_{W-}(s)$ we use Lemma 3 and the technique of spectral factorization [7]. We may write equation (3.10) as

$$L_{W-}(s) = x(s)\, L_{W-}(s) + Y(s) - z(s)$$

or

$$\{1 - x(s))\}\, L_{W-}(s) = Y(s) - z(s) \quad \rightarrow (3.13)$$

In order to use the spectral factorization technique we need to express

$$1 - x(s) = \frac{\varphi^+(s)}{\varphi^-(s)} \quad \rightarrow (3.14)$$

where $\varphi^+(s)$ is analytic in the right half plane and $\varphi^-(s)$ is analytic in the left half plane. From equation (3.10)

$$x(s) = \left(\frac{\mu}{\mu+s}\right)^n \prod_{i=1}^{n} L_{a_i}(-s) \quad \rightarrow (3.14a)$$

As $L_{a_i}(s)$ is the Laplace transform of a density function $a_i(t)$ and $\prod_{i=1}^{n} L_{a_i}(s)$ is the Laplace transform of the sum of n random variables coming from densities $a_1(t)$, $a_2(t)$ ... $a_n(t)$. Thus, $\prod_{i=1}^{n} L_{a_i}(s)$ is a Laplace transform of a valid density function. By substituting

$$\frac{\mu+s}{\mu} = z$$

we may express the equation

$$1 - x(s) = 0 \qquad \rightarrow (3.14b)$$

in the form of (3.11b). From Lemma 5 this equation has n roots. Let the negative of these roots be ζ_i, $i = 1,2,\ldots,n$. We note that $s = 0$ is also a root of this equation. Using these roots we may write $\varphi^+(s)$ and $\varphi^-(s)$ as

$$\varphi^+(s) = \frac{s \prod_{i=1}^{n} (s+\zeta_i)}{(s+\mu)^n} \qquad \rightarrow (3.15)$$

and

$$\varphi^-(s) = \frac{s \prod_{i=1}^{n} (s+\zeta_i)}{(s+\mu)^n(1-x(s))} \qquad \rightarrow (3.16)$$

Clearly, $\varphi^+(s)$ and $\varphi^-(s)$ satisfy the necessary conditions for spectral factoriztion [7]. From equations (3.13) and (3.14) we write

$$L_{W-}(s) \quad \varphi^+(s) = \varphi^-(s) \, [Y(s) - z(s)] \qquad \rightarrow (3.17)$$

From Lemma 2

$$\varphi^-(s) \, [Y(s) - z(s)] = \Phi^+(s) + \Phi^-(s) \qquad \rightarrow (3.18)$$

where

$$\Phi^+(s) = \sum_{i=1}^{n} \frac{c_i}{(\mu+s)^i} \qquad \rightarrow (3.19)$$

From equations (3.17) and (3.18)

$$L_{W-}(s)\varphi^+(s) - \Phi^+(s) = \Phi^-(s)$$

A straightforward use of the analytic continuation arguments and Liouville's Theorem [7] implies

$$L_{W-}(s)\varphi^+(s) - \Phi^+(s) = \Phi^-(s) = K$$

where K is a constant.

Therefore,

$$L_{w-}(s) = \frac{K - \Phi^+(s)}{\varphi^+(s)}$$

Substituting from equations (3.15) and (3.19), we have

$$L_{w-}(s) = \frac{(s+\mu)^n}{s \prod_{i=1}^{n} (s+\zeta_i)} \left[K + \sum_{i=1}^{n} \frac{c_i}{(\mu+s)^i}\right] \quad \rightarrow (3.20)$$

To evaluate K we note that $\lim_{t \to \infty} E_{w-}(t) = 0$

Hence,

$$\lim_{s \to 0} sL_{w-}(s) = 0 = \lim_{s \to 0} \frac{(s+\mu)^n}{\prod_{i=1}^{n} (s+\zeta_i)} \left[K + \sum_{i=1}^{n} \frac{c_i}{(\mu+s)^i}\right]$$

This implies that

$$K = - \sum_{i=1}^{n} \frac{c_i}{\mu^i} \quad \rightarrow (3.21)$$

Using equations (3.21) in (3.20) and simplifying we get

$$L_{w-}(s) = \frac{d_{n-1}s^{n-1} + d_{n-2}s^{n-2} + \ldots + d_0}{\prod_{i=1}^{n} (s+\zeta_i)}$$

where d_i's are constants.

Or, $$L_{w-}(s) = \sum_{i=1}^{n} \left(\frac{\alpha_i}{s+\zeta_i}\right)$$

where α_i's are constants.

This completes the proof. ■

Theorem 3 - A $G^n|M|1$ queue has a steady state solution if $\mu\beta > n$ where

$$\beta = \sum_{i=1}^{n} \int_0^\infty t\, a_i(t)dt$$

Proof - From Lemma 5 we observe that when $\mu\beta \leq n$ one of the n roots of the equation (3.11b) becomes $\delta_n = 1$. This root corrosponds to $\zeta_n = 0$. Thus from Theorem 2 $L_{w-}(s)$ has a term of the form $\frac{\alpha}{s}$.

As a consequence $\lim_{s \to 0} s\, L_{W-}(s)$, and hence $\lim_{t \to \infty} E_{W-}(t)$ is finite. This cannot be true for a steady state solution of the queue. Therefore the queue reaches a steady state if and only if $\mu\beta > n$ and hence the theorem. ■

The existence of a steady state has been established in Theorem 3 under the condition $\mu\beta > n$. As β is simply the expected length of time for n arrivals, this condition is no different from the standard steady state condition of $\rho < 1$.

A $G|M|1$ queue corresponds to a $G^n|M|1$ queue with $n=1$, since the solutions given above for the $G^n|M|1$ queue correspond to the well known solutions of $G|M|1$ queue when we take $n = 1$.

The form of $L_{W-}(s)$ given by equation (3.12) implies that in steady state $E_{W-}(t)$ can be expressed as a sum of n exponential terms where the exponent coefficients are determined from the roots of the equation (3.14b). From the form of x(s) as given in equation (3.14a), the roots of equation (3.14b) do not depend on the order of a_i's in the sequence. The initial value, $E_{W-}(0)$ does depend on the sequence, however.

To obtain the steady state solution of a $G^n|M|1$ system, one approach is to apply the spectral factorization technique used in the proof of Theorem 2 to determine the coefficients α_i's. The coefficients ζ_i's are obtained as roots of equation (3.14a) using the results of Lemma 5. An alternate approach to obtaining the coefficients α_j's is to assume

$$E_{W-_1}(t) = \sum_{j=1}^{n} \alpha_j e^{-\zeta_j t}, \qquad \to (3.22)$$

and substitute this in equations (3.2) and (3.3) . Also,

$$E_{W-_1}(0) = \sum_{j=1}^{n} \alpha_j$$

and

$$E_{W+_1}(0) = \frac{1}{\mu} + \sum_{j=1}^{n} \alpha_j$$

Thus from equation (3.2) we have

$$E_{w+_1}(t) = \left\{\frac{1}{\mu} + \sum_{j=1}^{n} \alpha_j\right] e^{-\mu t} + \mu e^{-\mu t} \circledast \left[\sum_{j=1}^{n} \alpha_j e^{-\zeta_j t}\right]$$

$$= \left\{\frac{1}{\mu} + \sum_{j=1}^{n} \frac{\alpha_j \zeta_j}{\zeta_j - \mu}\right] e^{-\mu t} + \sum_{j=1}^{n} \frac{\alpha_j \mu}{\mu - \zeta_j} e^{-\zeta_j t} \qquad \rightarrow (3.23)$$

Using equation (3.3) along with equation (3.23) we get

$$E_{w-_2}(t) = \left\{\frac{1}{\mu} + \sum_{j=1}^{n} \frac{\alpha_j \zeta_j}{\zeta_j - \mu}\right\} e^{-\mu t} \cdot L_{a_2}(\mu)$$

$$+ \sum_{j=1}^{n} \alpha_j \left(\frac{\mu}{\mu - \zeta_j}\right) L_{a_2}(\zeta_j) e^{-\zeta_j t} \qquad \rightarrow (3.24)$$

As the form of $E_{w-}(t)$ remains the same for each step in the cycle, in equation (3.24) we must have

$$\frac{1}{\mu} + \sum_{j=1}^{n} \frac{\zeta_j \alpha_j}{\zeta_j - \mu} = 0 \qquad \rightarrow (3.25)$$

Then

$$E_{w-_2}(t) = \sum_{j=1}^{n} \alpha_j \left(\frac{\mu}{\mu - \zeta_j}\right) L_{a_2}(\zeta_j) e^{-\zeta_j t} \qquad \rightarrow (3.26)$$

Similarly,

$$E_{w-_3}(t) = \sum_{j=1}^{n} \alpha_j \left(\frac{\mu}{\mu - \zeta_j}\right)^2 L_{a_2}(\zeta_j) L_{a_3}(\zeta_j) e^{-\zeta_j t}$$

and

$$\frac{1}{\mu} + \sum_{j=1}^{n} \alpha_j \left(\frac{\zeta_j}{\zeta_j - \mu}\right)\left(\frac{\mu}{\mu - \zeta_j}\right) L_{a_2}(\zeta_j) = 0 \qquad \rightarrow (3.27)$$

After the n^{th} step we have

$$E_{w-_1}(t) = E_{w-_{n+1}}(t) = \sum_{j=1}^{n} \alpha_j \left(\frac{\mu}{\mu-\zeta_j}\right)^n \left\{\prod_{k=1}^{n} L_{a_k}(\zeta_j)\right\} e^{-\zeta_j t} \qquad \rightarrow (3.28)$$

yielding n linear equations of the type (3.27).

Clearly, the equality of $E_{w-_1}(t)$ and $E_{w-_{n+1}}(t)$ holds where

$$\left(\frac{\mu}{\mu-\zeta_j}\right)^n \left\{\prod_{k=1}^{n} L_{a_k}(\zeta_j)\right\} = 1 \qquad \text{for } j = 1,2\ldots n \qquad \rightarrow (3.29)$$

The ζ_i's obtained as negative roots of equation (3.14b) satisfy equation (3.29). The values of α_i's are obtained by solving n linear equations of the form (3.25), (3.27), etc. A solution for a $G^n|M|1$ queue may, therefore, be obtained by finding n roots of equation (3.14b) and solving a set of n linear equations. This approach is simpler than the spectral factorization technique. In the next section we present an example of obtaining the steady state solution for a $G^n|M|1$ queue.

IV. APPLICATION

A situation in which $G^n/M/1$ queues occur is when an arrival stream is split among k servers using a deterministic sequence, the arrivals seen by each server form a cyclic pattern. For example consider the 2 server system shown in Figure 3. The arrivals at box D are

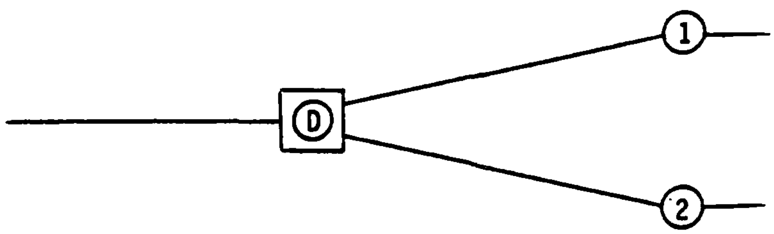


Figure 3

sent to the two servers using a sequence of length 3

1 1 2

Let the arrivals to D be independent with a general distribution F(t).
Let us denote the i-fold convolution, $F^{(i)}(t)$ by

$$F^{(i)}(t) = \overbrace{F(t) \circledast F(t) \circledast \ldots \circledast F(t)}^{i \text{ times}}$$

The arrivals seen by server 1 follow a cyclic pattern

$$F^{(1)}(t),\ F^{(2)}(t),$$

Similarly, the arrivals at servers 2 follow general arrivals from

$$F^{(3)}(t),$$

Queue 1 is of the $G^2|M|1$ type and can be analysed using the results of this paper.

In particular, let us consider the arrivals to D be Poission with rate λ , i.e.,

$$F(t) = 1 - e^{-\lambda t}$$

Let the service rates of the two servers be μ_1 and μ_2. For server 1, $F_1(t)$ is the same as $F(t)$ and $F_2(t)$ is given by

$$F(t) \circledast F(t)$$

Thus, $L_{a_1}(s)$ and $L_{a_2}(s)$ are $\frac{\lambda}{\lambda+s}$ and $\left(\frac{\lambda}{\lambda+s}\right)^2$, respectively. Equation (3.14b) reduces to

$$\left(\frac{\mu_1}{\mu_1+s}\right)^2\left(\frac{\lambda}{\lambda-s}\right)^3 = 1 \qquad \rightarrow (4.1)$$

Also, we have the two linear equations

$$\frac{1}{\mu_1} + \frac{\zeta_1\mu_1}{\zeta_1-\mu_1} + \frac{\zeta_2\zeta_2}{(\zeta_2-\mu_1)} = 0 \qquad \rightarrow (4.2)$$

and

$$\frac{1}{\mu_1} + \frac{\zeta_1\alpha_1}{\zeta_1-\mu_1} \cdot \frac{\mu_1}{\mu_1-\zeta_1} \cdot \left(\frac{\lambda}{\lambda+\zeta_1}\right)^2 + \frac{\zeta_2\alpha_2}{(\zeta_2-\mu_1)} \frac{\mu_1}{(\mu_1-\zeta_2)} \left(\frac{\lambda}{\lambda+\zeta_2}\right)^2 = 0$$

$$\rightarrow (4.3)$$

As a numerical example let $\mu_1 = 0{\cdot}2$ and $\lambda = 0{\cdot}1$.

In this case we get

$$\zeta_1 = 0{\cdot}14913$$

$$\zeta_2 = 0{\cdot}23295$$

$$\alpha_1 = 1{\cdot}41565$$

$$\alpha_2 = -0{\cdot}10789$$

Thus $E_{W-}(t) = 1{\cdot}41565\, e^{-0.14913t} - 0.1079\, e^{-0.23295t}$

From this we note that,

$$E_{W-}(0) = 1{\cdot}30776$$

$$E_{W+}(0) = 6{\cdot}30776$$

The before and after arrival values of the expected virtual waiting time, after the next arrival, are 2·4556 and 7·4556, respectively.

A solution for this example obtained by an iterative solution of the transient equations has been reported in [4]. The solution obtained here is exactly the same as the one reported in [4].

For this example, the interarrival times seen by server 2 have the same three stage hypo-exponential distribution. The expected waiting time seen by an arrival routed to this server by the above sequence can be obtained by using the standard solution to G|M|1 queue. For $\mu = 0.1$, the expected waiting time is 11.9403. Thus, the expected waiting time for an arrival to these two servers using the deterministic routing is 8.5679. If the routing were done at random with probability 2/3 for server 1 and 1/3 for server 2 the expected waiting time would have been 10.00. It is interesting to note that in this case a deterministic decision rule does better than a probabilitic rule.

V. SUMMARY

In this paper the steady state behavior of an exponential, first-in-first-out queue with cyclic arrivals has been analyzed. This arrival structure is a generalization of $G|M|1$ queues. The well known result of $G|M|1$ queues that the waiting time is exponentially distributed has been generalized to the $G^n|M|1$ queue. This generalization yields an easier technique for solving a given $G^n|M|1$ queue. An example to show the use of the results has been presented, as well.

VI. REFERENCES

[1] Cohen, J.W., "The Single Server Queue", North-Holland, Amsterdam, 1969.

[2] Lindley, D.V., "The Theory of Queues with a Single Server", Camb. Phil. Soc., Vol. 48, 1952, pp. 227-289.

[3] Kleinrock, L., "Queueing Systems: Vol. 1", Wiely, New York, 1976.

[4] Agrawala, Ashok K. and Tripathi, Satish K., "Transient Solution of the Virtual Waiting Time of a Single Server Queue and its Applications", TR-806, Computer Science, University of Maryland, August 1979. [To appear in International Journal of Information Sciences, Vol. 20, No. 3, April 1980.]

[5] Papoulis, A., "Probability, Random Variables, and Stochastic Processes", McGraw-Hill, New York, 1965.

[6] Takacs, Lajos, "Introduction to the Theory of Queues", Oxford University Press, New York, 1962.

[7] Carrier, G.F., Krook, M., and Pearson, C.E., "Functions of a Complex Variable", McGraw-Hill, New York, 1966.